STUDIA UBB PHILOLOGIA, LXII, 2, 2017, p. 167 - 182
(RECOMMENDED CITATION)
DOI:10.24193/subbphilo.2017.2.12

DEVELOPING AN ANALYSIS GRID FOR B1/ B2 READING TASKS AND ITEMS

ADRIANA TODEA[1]

ABSTRACT. ***Developing an Analysis Grid for B1/ B2 Reading Tasks and Items.*** A test analysis grid is a useful tool in the process of test design/ writing and test validation as its development can help both set up guidelines for test writers and provide a means of checking whether these standards are met, not only by test tasks globally, but by test items individually. Assessing reading abilities in a foreign language on CEFR standards requires that test analysis grids adapt to CEFR descriptors. The question is whether the CEFR model of language proficiency may prove particularly successful at building quantifiable scales of selected assessment criteria that can be used to work out scores of test tasks, and test items individually, which can then translate into CEFR levels. The purpose of this research is putting in place an effective algorithm to measure, on the CEFR scale, the various input text characteristics, item characteristics, and problem-solving strategies of reading test tasks in interaction. The present paper focuses specifically on the development of a test analysis grid that can provide an accurate tool to measure/ validate B1 and B2 reading test tasks and items of the Romanian language tests developed at the Babeş-Bolyai University by the Romanian Language, Culture and Civilization Department.

Key words: *test design, test validation, input text, rubric, task, item, expected response, reading strategies, problem-solving strategies.*

REZUMAT. ***Conceptul şi structura grilei de analiză pentru testele de înţelegere-citire de nivel B1 şi B2.*** O grilă de analiză a cerinţelor unui test de competenţă lingvistică este un instrument util în procesele de concepţie, structurare, creare şi validare a testului, ea putând fi folosită atât ca un mijloc de îndrumare a creatorilor de teste prin setul propriu de standarde şi specificaţii, cât şi ca un instrument de evaluare a cerinţelor şi itemilor creaţi de către experţi în procesul de validare. Evaluarea competenţei de înţelegere-citire pe baza standardelor CECRL (*Cadrul European Comun de Referinţă pentru Limbi*) face necesară adaptarea grilei de analiză la descriptorii cadrului european, în măsura în care acest model european standardizat al competenţei lingvistice într-o limbă străină poate susţine cu succes dezvoltarea unui sistem de evaluare cuantificabil

[1] Adriana Todea, PhD, Babeş-Bolyai University, atodea@yahoo.com

şi discriminatoriu între nivelurile de competenţă lingvistică. Scopul acestei cercetări este de a concepe un algoritm eficient de măsurare, pe scala CECRL, a caracteristicilor componentelor unui test de înţelegere-citire (textul sursă, textul itemului/ întrebării şi strategiile de procesare a informaţiei pentru identificarea răspunsului fiecărei întrebări) în interacţiune. În mod specific, acest demers are ca finalitate conceperea unei grile de analiză a testelor de înţelegere-citire dezvoltate de către Departamentul de Limbă, cultură şi civilizaţie românească a Universităţii Babeş-Bolyai în cadrul evaluării la limba română a studenţilor străini înscrişi în anul pregătitor.

Cuvinte cheie: *designul testului, procesul de validare, textul sursă, explicaţia şi instrucţiunile cerinţei, cerinţa, itemul, răspunsul preconizat, strategiile de lectură, strategiile de rezolvare a problemei.*

The assessment of language competence in relation to communicative skills, such as reading, is inherently dependent on correlating language knowledge thresholds with effective or successful task performance. Such an endeavor requires clear, contrastive and positive qualitative descriptions of what task success means for each grade on the assessment scale. The Common European Framework for Languages (2000), from this point forward CEFR, offers a language-independent assessment scale of language proficiency, complete with full descriptors of what effective performance means in relation to communicative skills, strategies, language knowledge and targeted content. As such, CEFR has been selected as the assessment theoretical framework for the L2 Romanian language tests developed at the Babes-Bolyai University by the Romanian Language, Culture, and Civilization Department, in order to bring into agreement the Romanian language assessment of L2 Romanian students with European standards of language assessment and, consequently, increase the national and international acceptability of the language certificates issued.

One of the first stages in the process of test validation is to establish test specifications in correlation with CEFR scales and descriptors as a basis for test design. Developing test analysis grids for the assessment of specific communicative skills is a constructive stage with multiple effects. Firstly, a test analysis grid constitutes a coherent and professional blueprint for test developers and designers by drawing clear dynamic correlations between test profile, test content, language thresholds, information processing operations, test solving strategies, background knowledge, etc, whose analysis results in an in-depth understanding of the detailed processes underlying the expected response, and, therefore, can support the development of a reliable,

documented system of discrete measurement. Secondly, the expert test validation and test verification processes can adapt the test analysis grid to a system of quantifiable criteria for measuring the conformity or non-conformity of tasks and items to test specifications. Ultimately, a reading test analysis grid explains how the interaction between input text features and the corresponding item features provides an effective and reliable framework for evaluating processes that occur naturally in authentic reading experiences.

Input text features

The features of the input text constitute the first criterion to decide the appropriateness of the reading task to the proficiency level it targets as a testing device. A successful reading experience relies on both knowledge of language (vocabulary, grammar) and reading skills, and therefore, testing reading should consider the appropriate thresholds for both language knowledge and reading abilities. As the L2 Romanian language tests are designed for adult learners who, at the very least, have completed a highschool education and are undergoing a Romanian language preparatory year in order to enroll in Romanian- medium universities, concerns relating to the reading skills threshold can be safely dropped, as candidates are expected to transfer their fully developed reading skills from L1 to L2. As such, candidates are not expected to have difficulties in dealing successfully with any type of text whose basic genre features and generic linguistic conventions are part of any L1 secondary education curricula. Furthermore, candidates are expected to make use of the reading strategies they have already developed in L1 in solving L2 reading tasks, such as skimming, scanning, or using context to understand an unknown word, in other words, to be fluent L1 readers. Thus metalinguistic and metacognition thresholds are not independent factors to be considered in test design, unless they are significantly impacted by language knowledge thresholds and task specificity.

Overall comprehension: skimming

As a form of expeditious reading at the global level of a text, skimming is a selective reading strategy that targets an overall comprehension and, as such, it is the means by which readers make first contact with the text. Skimming relies on our ability to recognize and understand the significance of text layout, text organization, paragraph organization and genre conventions. Successful skimming should provide an overall understanding of what the text is about and how the information is organized, that is, allow the reader to build a mental *framework to accommodate the whole text* (Weir, Huizhong & Yan, 2000: 25).

Text organization

CEFR gives one important description as to the B1 appropriateness of level of text organization in two of its reading scales:

B1: "Can read *straightforward* factual texts on subjects related to his/ her field and interest with a satisfactory level of comprehension." (Overall reading comprehension scale)

B1: "Can identify the main conclusions in *clearly signaled* argumentative texts. Can recognize significant points in *straightforward* newspaper articles on familiar subjects." (Reading for information and argument scale)[2]

Straightforward … and clearly signaled … text are descriptions of text clarity, of which text organization is a key component. Taking into account that in the same scales CEFR provides contrastive text qualifiers such as *simple* (A2) and *complex* (C1-C2), the required straightforward clarity of text organization is a property of B2 as well, by virtue of its hierarchical inclusiveness of B1). Thus, comprehensible skimming at B1/B2 would require a *standard or conventional* organization of the text according to genre conventions and text layout.

Skimming and lexis

Skimming relies primarily on language knowledge, though. The overall understanding of the text depends, first and foremost, on its lexis. The adequate comprehension of any text relies on the understanding of 95% of the words in the text (Alderson, 2005: 35), with the meaning of the remaining 5% being worked out from context. Skimming, though, does not mean the full, detailed, careful reading of the whole text, but the mostly expeditious, occasionally careful, selective reading of text parts such as: title, subtitles, headings, introduction, conclusion and topical sentences (Weir, Huizhong & Yan, 2000: 40). Therefore, for establishing a general understanding of a text, the lexical and morpho-syntactic difficulty appropriate for the target level should not be judged in terms of overall percentages, but evaluated specifically in the parts of the text targeted by skimming.

Text content: complexity and familiarity of topic

The appropriateness of the input text should be judged as well in terms of its content: comprehensibility of content relies on a balance between the complexity of information and the familiarity of the subject matter. According to Weir, Huizhong & Yan (2000: 25-28), activating background knowledge acts like a

[2] CEFR, pp. 69-70, emphasis added.

coping mechanism for readers with poor language knowledge, and improves the comprehension of highly specialized texts for readers with good language knowledge. Similarly, particularly for B1 and B2 levels, CEFR balances comprehension of complex content against its familiarity, whereas the A2 and C1 descriptors lack such a degree of complexity- degree of familiarity correlation. In other words, A2 comprehension is restricted to texts that lack any degree of complexity (*short simple texts*), and C1 comprehension lacks any degree of familiarity caveat (*whether or not they relate to his/ her own area of speciality)*.

A2: "Can understand *short simple texts* on familiar matters of a concrete type which consists of high frequency everyday or job-related language" (Overall reading comprehension scale)

B1: "Can read *straightforward factual* texts on *subjects related to his/ her field and interest* with a satisfactory level of comprehension." (Overall reading comprehension scale)

B2: "Can understand articles and reports concerned *with contemporary problems* in which the writers adopt *particular stances or viewpoints*." (Reading for information and argument scale)

B2: "Can obtain information, ideas, and opinions *from highly specialized sources within his/ her field*. Can understand *specialized articles outside his/ her field*, provided he/ she can *use a dictionary* occasionally to confirm his/ her interpretation of *terminology*." (Reading for information and argument scale)

C1:" Can understand in detail a *wide range of lengthy complex texts* likely to be encountered in social, professional or academic life, identifying points of detail including *attitudes* and *implied as well as stated opinions*." (Reading for information and argument scale)

C1: "Can understand in detail *lengthy, complex texts, whether or not they relate to his/ her own area of speciality*, provided he/ she can reread difficult sections."[3] (Overall reading comprehension scale)

B1 comprehension allows complexity at the level of factual, concrete information provided the content is familiar, whereas B2 allows comprehension of both factual and abstract information (*specialized articles, articles and reports concerned with contemporary problems, highly specialized sources*) by degrees, as follows: partial spontaneous comprehension (*use a dictionary*) of complex texts (*specialized articles*) *outside his/ her field*, but full comprehension of *information, ideas, and opinions* on familiar topics *from highly specialized sources* or on topics of general interest (*contemporary problems*). It is at these two particular levels, according to CEFR, that background knowledge or subject matter knowledge can facilitate comprehension provided the threshold language knowledge is reached

[3] CEFR, pp. 69-70, emphasis added.

(Alderson, 2005: 102-106). According to the CEFR descriptors above, the language knowledge threshold for spontaneous unaided successful comprehension of unfamiliar factual and abstract content is C1. If assisted comprehension (*use of dictionary*) is not a test specification, then text selection in B1/B2 reading tests should be restricted to those of familiar content (either specialized or of general interest).

To sum up, a B1/B2 test analysis content scale should present increasing content complexity, but only on familiar topics.

Lexis

Providing an appropriate scale for levels of knowledge of grammar and vocabulary required to support successful comprehension is not one of the CEFR's strengths. Both the General linguistic range and Vocabulary range scales are language production and not language comprehension oriented. The Vocabulary range scale quoted below describes lexical competence in terms of successful linguistic expression that balances, in particular for A2, B1 and B2, degrees of complexity against degrees of familiarity of various kinds of communicative interactions.

A2: "Has sufficient vocabulary for coping with simple survival needs; for the *expression* of basic communicative needs; to conduct routine, everyday transactions involving familiar situations and topics."

B1: "Has sufficient vocabulary to *express* himself/ herself with some circumlocutions on most topics pertinent to his everyday life such as family, hobbies and interest, work, travel and current events."

B2: "*Has a good range of vocabulary* for matters connected to his field and most general topics. Can *vary formulation* to *avoid frequent repetition*, but lexical gaps can still cause *hesitation* and *circumlocution*."

C1: "Has a good command of a *broad lexical repertoire* allowing gaps to be readily *overcome* with *circumlocutions*; little *obvious searching for expressions* or *avoidance strategies*. Good command of idiomatic expressions and colloquialisms."[4] (Vocabulary range scale)

As linguistic comprehension constantly outpaces production, knowledge, the above scale is not particularly useful in designing a language level scale appropriate for testing reading skills, aside from suggesting that we might give a description of active reading vocabulary based on the content complexity descriptions in the reading scales themselves.

The Overall reading comprehension scale provides some clearly stated vocabulary comprehension descriptions, though not consistently, as follows:

[4] CEFR, p.112, emphasis added.

A2: "Can understand short simple texts on familiar matters of a concrete type which consists of *high frequency everyday or job-related language*"

B1: "Can read straightforward factual texts on subjects related to his/ her field and interest with a satisfactory level of comprehension."

B2: "Can read with a large degree of independence, adapting style and speed of reading to different texts and purposes, and using appropriate reference sources selectively. Has a *broad active reading vocabulary*, but may experience some *difficulty with low-frequency idioms*."

C1: "Can understand in detail lengthy, complex texts, whether or not they relate to his/ her own area of speciality, provided he/ she can reread difficult sections."[5] (Overall reading comprehension scale)

Level A2 and B2 descriptors are the only ones that contain specific language/ vocabulary descriptions, that is level A2 understands *high frequency everyday or job-related language*, whereas B2 has a *broad active reading vocabulary* and, by inference, has no *difficulty* understanding high-*frequency idioms*. Significantly, the B2 description of lexical comprehension correlates with the C1 Vocabulary range descriptor quoted above:

-- *broad active reading vocabulary* vs. *broad lexical repertoire*;

-- no *difficulty* understanding high *frequency idioms* vs. *good command of idiomatic expressions and colloquialisms*;

and not with its corresponding B2 Vocabulary range scale descriptor: *Has a good range of vocabulary for matters connected to his field and most general topics*.

By extrapolation, A2 and B2 reading comprehension descriptors can be used as reference points for assigning comprehension vocabulary values in between the two, B1, and above B2, C1, in correlation with text content descriptions. Hence, we can construct a *reading vocabulary scale* as follows:

A2: Can understand the *high frequency* vocabulary of texts consisting of *simple familiar* factual information;

B1: Can understand the *frequent* vocabulary of texts consisting of c*omplex*, but *familiar*, facts;

B2: Can understand the *broad vocabulary range/ low frequency vocabulary* and the *high frequency idioms* of texts consisting of *complex* facts, ideas and opinions on *familiar* topics;

C1: Can understand the *extended vocabulary range* and the *low frequency idioms* of texts consisting of *complex* facts, ideas and opinions on *unfamiliar* topics.

In this case, the C1 vocabulary descriptor can logically expand the B2 *broad range* qualifier to the C1 *extended range* and the B2 *high-frequency idioms* to the C1 *low-frequency idioms*. On the other hand, the A2 *high frequency* qualifier

[5] CEFR, p. 69, emphasis added.

can generate two more scale values as *frequent* B1 and *low- frequency* B2. But B2's low frequency vocabulary cannot include specialized vocabulary *outside his/ her field*, which, according to the B2 descriptor in the Reading for information and argument scale, would require the use of a dictionary. Therefore, reading vocabulary descriptors need to combine, correlate and generate degrees of

- *language use* (vocabulary frequency scale);
- *language domains* (general vs. idiomatic, general vs. specialized);
- *background knowledge* (familiarity);
- *content complexity* (abstract vs. factual, information complexity).

	A2	B1	B2	C1
words and expressions	high frequency	frequent	some low frequency	many low-frequency
idioms	--	--	high frequency	low frequency
vocabulary range	limited	good	broad	extended
concrete information	concrete	concrete	mostly concrete	some concrete
abstract information	--	--	some abstract	mostly abstract
specialized information	--	low degree of specialization	highly specialized	highly specialized
information complexity	simple	complex	complex	complex
degree of familiarity	only familiar	mostly familiar	rather familiar	rather unfamiliar

Items as independent problem-solving tasks

Reading processes may very well vary depending on their specific purpose: pleasure, information, or learning. Reading tests, as components of L2 language competence examinations, though, would invariably turn the reading experience, in the evaluation process, into a problem-solving task, that would explicitly or implicitly require that the test-takers use their reading abilities to process information in very specific ways, such as

- follow instructions (invariably in every rubric text);
- identify gist, structure, main ideas, definitions, supporting details, points in a line of argumentation, specific details, relevant data or evidence, viewpoints, opinions, purpose, motivation, underlying theme or concept;
- infer attitudes; feelings, moods, purpose, motivation;

- compare data, facts, evidence, ideas and understand how they relate to one another as cause, effect, solution or purpose;
- draw logical inferences;
- reach a conclusion;
- evaluate attitudes, moods, overall purpose.

Testing reading abilities through specific operations of information processing like the ones above are nonetheless perfectly compatible with real life goals and motivations underlying general or specialized reading activities. In this respect, items designed on such operations have an appropriate degree of authenticity. There are two criteria that can narrow down the selection of an operation in item design: one is the language threshold necessary for successful task completion, the other is the integration of language knowledge, content and skill in ways that would make a language certificate particularly relevant to the needs, requirements and standards of stakeholders who may require general language proficiency or academic/ professional language proficiency certification, with various degrees of specialization. The B2 Romanian language test developed by the Romanian Language, Culture, and Civilization Department is specifically designed as an in-house exit test for students enrolled in the Romanian language preparatory year, and the certification is required for access to university programs with Romanian as language of instruction. As such, its general purpose is testing language knowledge and skills that are relevant to coping with university life and instruction. Although the B1 Romanian language test is administered to the same L2 Romanian students as a midterm exam, its general focus is language testing for communicative autonomy and social integration.

With respect to language thresholds that can support specific item operations, CEFR provides significant details in its Reading for information and argument scale, as follows:

A1: "Can *get an idea of the content* of simpler informational material and short simple descriptions, especially if there is visual support."

A2: "Can *identify specific information* in simpler written material he/she encounters such as letters, brochures and short newspaper articles describing events."

B1 "Can *recognise significant points* in straightforward newspaper articles on familiar subjects."

B1: "Can identify the main conclusions in clearly signalled argumentative texts. Can *recognise the line of argument* in the treatment of the issue presented, though *not necessarily in detail*."

B2: "Can *understand* articles and reports concerned with contemporary problems in which the writers adopt particular *stances* or *viewpoints*."

B2: "Can *obtain information, ideas and opinions* from highly specialised sources within his/her field. Can understand specialised articles outside his/her

field, provided he/she can occasionally use a dictionary to confirm his/her interpretation of terminology."

C1: "Can understand in detail a wide range of lengthy, complex text likely to be encountered in social, professional or academic life, *identifying finer points of detail* including *attitudes*, and *implied* as well as *stated opinions*."[6]

Providing that such information processes as described and highlighted above are understood in strict correlation with the degree of text complexity, and that the hierarchical organization of CEFR levels presupposes that the higher levels are inclusive of lower level abilities, we can correlate language thresholds to item problem-solving operations as follows:

	Problem solving operations	CEFR descriptors (excerpts)
C1	• identify underlying theme or concept; • infer attitudes; feelings, moods, purpose; motivation;	*identifying finer points of detail including attitudes, … implied … stated opinions*
B2	• identify supporting details, viewpoints, opinions, purpose, motivation; • compare and relate ideas; • evaluate attitudes, moods, overall purpose;	*understand ... stances … viewpoints; obtain information, ideas and opinions*
B1	• identify main ideas, points in a line of argumentation, relevant evidence; • compare and relate evidence; • reach a conclusion; • draw logical inferences;	*recognise significant points; recognise the line of argument …not necessarily in detail*
A2	• follow instructions, identify definitions, specific details, relevant data; • compare and relate data, facts	*identify specific information*
A1	• identify gist, structure	*get an idea of the content*

Thus, each reading test item becomes an assignment, which requires the use of one or more calibrated problem-solving operations, is to be undertaken mostly independently of others, and requires that test-takers identify and process relevant information from the input text, according to overt instructions stated in the rubric, and by means of procedures and processes which are part of the knowledge and skills under evaluation.

Phrasing the item

A successful completion of an item assignment depends, first and foremost, on a complete understanding of its requirements. Whereas some gaps in the comprehension of the input text may not hinder a successful

[6] CEFR, p.70, emphasis added.

completion of specific item tasks, the full comprehension of the rubric instructions and item text is absolutely essential. In this respect, the phrasing of the item and rubric instructions should be fully and easily comprehensible to the candidate, and therefore, should rigorously conform, in their entirety, to the descriptors of the language level targeted by the reading test or to the level immediately below, in both their lexical and morpho-syntactic components.

Scanning

Scanning is the reading strategy of selectively and expeditiously looking through a text in search for clues that will locate specific information, or the segment of the input text where a specific answer can be found. This strategy prevents the repeated, time consuming careful reading of the full text in order to complete the problem-solving assignment underlying each item. Scanning will pick key words (numbers, dates, names, words or phrases from the item text and find appropriate matches for them in the input text. The purpose of this reading strategy is not comprehension, but locating information (Weir, Huizhong & Yan, 2000: 25; 41). Effective scanning relies as well on the ability to capitalize on skimming, that is, to use the knowledge of textual features and organization to narrow down the areas of the text that need to be scanned.

Scanning can have different degrees of difficulty based on whether

- there is a perfect match in *form*, *morphological category*, and *meaning* between the key words selected from item text and their matching correspondents in the input segment (numbers, dates, and names);
- there is an imperfect match in *form*, but a perfect match in *morphological category*, and *meaning* between the key words selected from item text and their matching correspondents in the input segment (use of synonyms);
- there is an imperfect match in *form*, and *morphological category*, but a perfect match in *meaning* between the key words selected from item text and their matching correspondents in the input segment (a key part of input segment has been rephrased in the wording of the item; use of antonyms);
- there is an imperfect match in *form*, *morphological category* and *overt meaning* between the key words selected from item text and their matching correspondents in the input segment (the matching meaning has to be inferred from larger segments such as clauses, sentences, or paragraphs, or via related semantic fields).

Scanning features as a reading strategy/ ability in the CEFR Reading for orientation scale, although it is clear from the descriptions that the term is used inclusively for both scanning and search reading. Still, the CEFR descriptors below are relevant for establishing that both B1 and B2 test-takers are expected to be competent at it.

B1: "Can *find and understand relevant information* in everyday material, such as letters, brochures and short official documents."

B1: "Can *scan longer texts* in order to *locate desired information*, and gather information from different parts of a text, or from different texts in order to fulfil a specific task."

B2: "Can *scan quickly* through *long and complex texts, locating relevant details*."[7]

Thus, whatever scanning difficulties may occur at these levels, they are expected to arise from the size and language complexity of the targeted segments of the input text, rather than from incompetent scanning.

Search reading

Once the relevant segment is located in the input text, information is processed through search reading: a careful reading at a local level, within the target segment, but guided by the "predetermined topic" which results from the item's problem-solving assignment (Weir, Huizhong & Yan, 2000: 25, 38). The reading process at this level has to be thorough enough to generate an accurate representation of the segment. In this respect, the language level of the input segment has to be both relevant and accessible enough to the targeted level, although not exclusively so. Provided the segment is large enough, unknown words and expressions, or unfamiliar meanings could be worked out from the context. In other words, the phrasing of the input segment has to generate full accessibility to information by balancing any challenging language against input segment size and coherence.

As discussed in the scanning section, the CEFR Reading for orientation scale treats scanning and search reading inclusively, as a composite strategy/ ability measuring full reading comprehension at a local level. And, as stated there, the B1 and B2 descriptors assume full search reading competence in correlation to appropriate degrees of text complexity.

Conclusion

The present paper focuses on breaking down the dynamic interdependence between underlying strategies, operations, and linguistic knowledge in L2 reading experiences, in relation to text specificity. These analyses attempt to build reading task/ item feature scales, both evaluative and descriptive of general and/ or academic reading performance. Such scales could

[7] CEFR, p. 70, emphasis added.

constitute the basis for developing reading test specifications, test design, and expert test verification and validation. The annex below summarizes these results in what we consider to be a functional test analysis grid.

BIBLIOGRAPHY

Alderson, J. Charles *Assessing Reading*, Cambridge University Press, 2000

Amič Irena, Darja Šorjanc Braico, Zdravka Godunc, Nataša Elvira Jelenc, Ana Marija Muster, Branka Petek, and Tatjana Žlindra *The European Language Portfolio for Adults*, Slovenia, 2010, accessed February 10, 2014, elp.ecml.at/Portfolios/tabid/2370/PublicationID/59/language/en-GB/Default.aspx

Common European Framework of Reference for Languages: Learning, teaching, assessment, Council of Europe, 2000, accessed February 28, 2017, www.coe.int

The Dutch CEFR Grid Reading/ Listening, accessed November 28, 2016, lancaster.ac.uk/fss/projects/grid

Weir, J. Cyril, Yang Huizhong and Jin Yan, *Studies in Language Testing 12: An empirical investigation of the componentiality of L2 Reading in English for academic purposes,* Cambridge University Press, 2000

Annex

Test analysis grid for B1/B2 reading tasks and items in L2 Romanian language certification tests

☒= the variant is listed for contrastive purposes, but it is *not eligible*;
☐= *eligible* variant, providing it matches the required level, if specified.

Input text features

Text source: [*specify*]
Text genre: [*specify*]
Text type:
☐ mainly descriptive/ expository;
☐ mainly narrative;
☐ mainly argumentative.

Text consists of:
☒ mostly *complex* facts, ideas and opinions on *unfamiliar* topics;
☐ mostly *complex* facts, ideas and opinions on *familiar* topics *B2*;

☐ mostly c*omplex*, but *familiar*, facts *B1*;
☒ only *simple familiar* facts.

Text length:[*specify*]

Text and paragraph organization
☒ complex/ sophisticated/ unconventional;
☐ standard/ conventional;
☒ simplified/ rudimentary.

***Most* of the vocabulary in introduction, conclusion and topical sentences is within**

☒ an *extended* range of vocabulary, including *many low-frequency words* and *some low-frequency idioms*, which can convey both *complex concrete* and *abstract* information*, rather unfamiliar* and possibly *highly specialized*;
☐ a *broad* range of vocabulary, including *some low-frequency words* and *high-frequency idioms*, which can convey both *complex concrete* and *abstract* information, *rather familiar* and possibly *highly specialized*; B2
☐ a *good* range of *frequent* vocabulary which can convey *complex concrete* information, *mostly familiar* and possibly *slightly specialized*; B1
☒ *limited* range of *high frequency* vocabulary which can convey only *simple, concrete, familiar* information.

Item features

Item type
☐ multiple choice;
☐ True/ False;
☐ multiple matching.

Item's wording is within
☒ an *extended* range of vocabulary, including *many low-frequency words* and *some low-frequency idioms*, which can convey both *complex concrete* and *abstract* information*, rather unfamiliar* and possibly *highly specialized*;
☐ a *broad* range of vocabulary, including *some low-frequency words* and *high-frequency idioms*, which can convey both *complex concrete* and *abstract* information, *rather familiar* and possibly *highly specialized*; B2
☐ a *good* range of *frequent* vocabulary which can convey *complex concrete* information, *mostly familiar* and possibly *slightly specialized*; B1/B2
☐ *limited* range of *high frequency* vocabulary which can convey only *simple, concrete, familiar* information B1

Item's underlying problem-solving operations
☐ follow instructions;
identify ☐ gist of (part of) text;

☐ structure;
☐ main idea of paragraph;
☐ definition;
☐ supporting detail(s), B2
☐ point(s) in a line of argumentation;
☐ specific detail(s);
☐ relevant data or evidence;
☐ viewpoints; B2
☐ opinions; B2
☐ purpose; B2
☐ motivation; B2
☒ underlying theme or concept;

infer/ evaluate ☒ attitudes;
☒ feelings;
☒ moods;
☒ purpose;
☒ motivation;

compare ☐ data;
☐ facts;
☐ evidence;
☐ ideas; B2

identify information relationships: ☐ cause;
☐ effect;
☐ solution;
☐ purpose;
☐ draw logical inferences;
☐ reach a conclusion.

Scanning locates segments in the input text that match most item key words in

☐ form, morphological category, and meaning;
☐ morphological category, and meaning;
☐ overt meaning;
☐ related meaning;
☐ inferential meaning.

Input segment's vocabulary is mostly within

☒ an *extended* range of vocabulary, including *many low-frequency words* and *some low-frequency idioms*, which can convey both *complex concrete* and *abstract* information*, rather unfamiliar* and possibly *highly specialized*;

☐ a *broad* range of vocabulary, including *some low-frequency words* and *high-frequency idioms*, which can convey both *complex concrete* and *abstract* information, *rather familiar* and possibly *highly specialized*; B2

☐ a *good* range of *frequent* vocabulary which can convey *complex concrete* information, *mostly familiar* and possibly *slightly specialized*; B1
☒ *limited* range of *high frequency* vocabulary which can convey only *simple, concrete, familiar* information

Input segment size:
☒ one or more nonadjacent sentences from different paragraphs;
☐ one or more nonadjacent sentences within the same paragraph;
☐ one complex sentence/ more than one adjacent simple sentences;
☐ one clause/ simple sentence.